# 1

# The Nature and Problems of Platforms

Since this is a book about platform governance, it should start by saying what a platform is. As noted in the Introduction, the things that go by the name "platform" probably bear a family resemblance to one another rather than a rigid set of defining and distinguishing features. Accordingly, rather than attempt a formal definition, I propose to focus on the qualities that seem to make arrangements of humanity such as Facebook and Twitter so difficult to govern. Thus, I will use the term "platform" for a unified set of human-usable tools or services on the Internet that enables people not affiliated with the company that operates the tools to interact with one another, typically monetizes that interaction, and is characterized, generally, by the following properties (all of which are matters of degree rather than binaries):

1. It provides positive *network externalities*, that is, users benefit the more other users are present.
2. It provides strong *network affordances* for users to interact with distant (either in geographic or social terms) others.[1]
3. It operates at a *substantial scale* as a consequence of the incentives that its network character gives it.
4. It offers affordances for and is actually characterized by a *substantial diversity of uses and norms* as a consequence of the scale and network affordances (i.e., with lots of people who are different from one another, they find lots of things to do with the tool).

I'll fill out each of these ideas below, with some diversions and implications.[2]

It will often not be possible to decree, as a simple up or down determination, whether any single service is or is not a platform, but it will be possible to at least

[1] The notion of social distance captures the way that even a geographic platform might create more dense geographically bound social networks – if I join Nextdoor, for example, I can interact with many more of my neighbors than those whom I might ordinarily see.

[2] In the existing literature, my definition (such as it is) is perhaps closest to that of Srnicek (2016, 43–48). An earlier, related use of the term "platform" in the technology industry context centered on the capacity of other applications to run on a service, as with the idea of a mobile phone with an



https://doi.org/10.1017/9781108975438.002 Published online by Cambridge University Press

identify differences in platform-ness across different internet services with reference to the above properties.[3] For practical purposes, I will treat the creation or facilitation of significant behavioral novelty among users as a key indicator of the platform character of some service. The reason for this is pragmatic: While non-platform types of intermediaries also find themselves engaging in a certain amount of behavioral governance (e.g., payment processors like Stripe or the credit card networks regulate the sorts of goods and services their users may sell, such as by refusing to serve pornographers), the governance challenges of such intermediaries aren't nearly as interesting. They're basically just facilitating the same kinds of transactions that banks have facilitated for a very long time and can probably rely on similar regulatory strategies.[4]

Ultimately, I agree with Lobel (2016, 101–2) who identifies this as part of the nature of the domain: because the core of what platforms offer is a set of affordances

app store as a "platform" or the early market positioning of Facebook as a place to make applications, such as the infamous game "Farmville," available to its users. That alternative usage of the term is equivalent to the "computational" sense of platform described by Gillespie (2010, 349). This volume is not primarily concerned with "platforms" in the computational sense. Rather, the sense that seems most relevant to the current analysis is what Gillespie describes as "figurative," in which one uses the word "platform" to describe a collection of affordances allowing one to do something – in which we say, for example, that the president of the United States has a large platform in virtue of his or her capacity to reach a broad audience. Another helpful distinction in the existing literature is between "transaction" and "innovation" platforms (Cusumano, Gawer, and Yoffie 2019). Here, I focus entirely on "transaction" platforms.

[3] Another important question, with which this book cannot wrestle, is the distinction between the application layer and the infrastructure layer: When companies like Cloudflare decide to control the behavior of the companies (who themselves host users) who use their services (cf. Zittrain 2021), is this different from when Facebook does it? I am inclined to draw a fairly sharp line between the two in view of the fact that most infrastructure companies lack the network properties of what I am calling platforms, but this may be controversial. Relatedly, it is important to note that not all internet services that engage in the enterprise of "content moderation" are platforms, and the same service may function as a platform in one context but not in another. For example, workplace communication tools such as Slack and Basecamp may need to engage in content moderation, however, to the extent they tend to be deployed in closed universe settings on top of an existing physical network (such as within a team at a specific employer), they are unlikely to feature the same kind of diversity as something like Twitter. However, if they are deployed more broadly, for example, an open Slack instance for everyone in the world with a specific set of interests, they may function as platforms. Moreover, even in closed-universe contexts, companies conducting content moderation can experience problems similar to those described in this book, so the analysis herein may be of some use to them. For an amusing example: It turns out that at least one corporate communications service provider moderates the word "bone," doubtless because of its use as sexual slang and the problem of workplace sexual harassment. But that didn't work so well when a society of fossil-studying paleontologists subscribed to the service! Maria Cramer, "Paleontologists See Stars as Software Bleeps Scientific Terms," *New York Times*, October 18, 2020, www.nytimes.com/2020/10/18/science/paleontology-banned-words-convey.html.

[4] Even this attempt at line-drawing can break down: A company like Uber is a quintessential platform, but it is debatable whether it has created behavioral novelty to the degree of something like YouTube – perhaps it has, in view of the way it has blurred the lines between commercial and individual drivers, or perhaps its primary novelty is related to the scale of its operations. Oddly, these definitional boundaries can also blur in the other direction. Content creators who use YouTube can also treat it like the Uber sort of "gig economy" arrangement to the extent they rely on payments from YouTube's monetization program (Caplan and Gillespie 2020, 2).

to connect users, and because those affordances are compatible with a huge array of economic (and noneconomic or semi-economic) activities, in the context of cultural as well as economic incentives for continual innovation, platforms will necessarily resist firm definitional line-drawing.[5]

## 1.1 THE CORE FEATURES OF PLATFORMS

### 1.1.1 *Positive Network Externalities*

The presence of positive network externalities is perhaps the most basic and well-understood feature of platforms. If I'm a member of a social media site, I typically get more value from that website to the extent that people I already know – or people I want to know – are also using it; likewise, if I'm either a seller or a buyer on transactional platforms like eBay, I get more value to the extent that there are more potential counterparties in the marketplace. This doesn't require much elaboration because it is so well known, but subsequent sections will fill out some of the related details.

### 1.1.2 *Interactive Network Affordances*

Platforms' core function from the user standpoint is to provide a system to enable a large number of people to engage in various kinds of interactions with one another. For convenience, we can call this central feature of a platform its organization around "interactive network affordances." Those interactions might be economic (eBay, Etsy), interpersonal (Facebook, Twitter, YouTube), or both (LinkedIn).

Much of the added value of any platform is its capacity to expand the effective scope of a network that a given user can access – at least potentially. If I have enough followers on Twitter, I can reach many more people with my hot takes on the Supreme Court than if I just speak to my friends. If I have something to sell, I can reach people all over the world on eBay rather than simply relying on my own city. Even in platforms nominally rooted in a user's own social network (Facebook) or in specific geographic locations (Airbnb), network expansion is a key affordance (interacting with friends of friends, discovering otherwise-unadvertised rental properties).[6]

[5] That being said, many of the challenges and techniques of platform governance are also challenges and techniques of intermediary governance more generally. The banking system provides numerous examples of, for example, the use by governments of private intermediaries to regulate their users (as with anti-money-laundering legislation that requires banks to report certain kinds of transactions to the government) – a relationship between government and company that might not be seen as meaningfully different from the ones that show up in major internet platform contexts with legislation such as the Digital Millennium Copyright Act.

[6] For a discussion of how one social network failed, at least within a specific context, in part because it could not deliver network novelty to its users, see Pearson and Moskal (2017, 122–26). For a useful early review of some of the economic concepts, see Evans and Schmalensee (2010).

"Expansion" is probably not the correct term in all cases, for human cognitive and time limitations continue to be a constraint on what we might call the effective size of a person's addressable network. I can only consider so many lodging options or regularly interact (except in a broadcast mode) with so many other people. Another key aspect of the network affordances that platforms tend to offer, rather than raw expansion, is customization of one's network (thanks in part to recommender algorithms, discussed below), which tends to allow people to find niches not otherwise available to them.[7]

This network niche-finding capacity can have dramatic consequences. Consider the phenomenon of meme stocks: A subreddit called "Wall Street Bets" became, for mysterious reasons, a kind of focal point for people who not only traded stock recreationally, but for those who were interested in the pure entertainment value of a kind of expressive stock buying – choosing, from some combination of nostalgia, nihilism, and perhaps a (successful) effort at market manipulation, to aggressively buy companies like GameStop. It turns out that lots of people were watching, and at least some of those watchers may have recognized the power of this group to affect pricing in otherwise lightly traded stocks. Elon Musk joined in.[8] Accordingly, lots of those watchers eventually bought in themselves on the anticipation or observation of price movements, and GameStop and other "meme stock" companies had massive rallies. There are lots of reasons for the rise of the meme stock, including the phenomenon of virality, discussed in detail below, but certainly one significant part of it is that Reddit created a global as well as publicly visible place for people who were both interested in stock trading and a little bit nihilistic (r/wallstreetbets bills itself as "Like 4chan found a Bloomberg Terminal").

Social media platforms are particularly novel because they tend to afford users the opportunity for multilateral interactions. That is, users may interact with a large number of other users at once. For example, on social media, users in their capacity as content producers may potentially reach an unlimited number of others both sequentially (as with "retweeting") and simultaneously (as when a user with a lot of followers has broad reach for a single piece of content). Similarly, a user in his or her capacity as content consumer may consume content from a large number of others, as by following many people on Twitter. This quality of social media derives in part from the fact that (unlike the sale of physical products by transactional platforms) the things (content) that users supply to one another are non-rivalrous, and hence suitable for resharing (and with it distinctive problems like virality) or for bulk

[7] For example: I own far too many Apple Watch bands (there are lots of them; it's kind of a fashion accessory). It turns out there are several subreddits and a Discord chat that allow me to talk to other people who also own far too many of them, trade extras, find deals off the (outrageous) face price, and so forth – reminiscent of the baseball card conventions of my youth in the 1980s and 1990s, except permanently available everywhere. (You might reasonably question the social value of this.)

[8] By posting a link to r/WallStreetBets in a tweet with the text "Gamestonk!!" (https://twitter.com/elonmusk/status/1354174279894642703, January 26, 2021).

https://doi.org/10.1017/9781108975438.002 Published online by Cambridge University Press

consumption and second-order interaction (such as Facebook or YouTube comments). Indeed, we can identify social-media-like phenomena (including potential governance problems) in seemingly nonsocial platforms by observing that users primarily interact with one another by exchanging non-rivalrous content. For example, the open-source software code-sharing website GitHub, though very different from what we think of as "social media," could display similar network dynamics and cause similar social problems as the social media platforms.[9]

### 1.1.3 *Skim as a Revenue Model and Incentives to Scale (and Maybe Concentration)*

Commercial platforms primarily make money by skimming from the top of the activity that they facilitate. This is most obvious in the case of platforms oriented around economic activity, which typically take an explicit commission. But it is also true of platforms oriented around social activity, which skim some of the attention of their users off of the top and sell it. For the most obvious (and, to a professor who sometimes tries to show videos in class, deeply annoying) example, YouTube requires its users to devote some of their attention not to content created by their fellow users but to advertisements.

The skim-based revenue model contributes to strong platform incentives to scale – to increase users – as revenue is increasing in number of active users.[10] Moreover, it also gives platforms economic incentives to promote increased activity by existing users ("engagement" in the social media context).

The prominence of positive network externalities in the incentive structure surrounding platforms implies that they tend to scale organically: for both companies and users, the benefit received from a platform is increasing (only up to a point for users – see Section 1.1.4 for recommender algorithms) with the size of the platform's userbase, so platform growth tends to be self-reinforcing. Successful social media platforms in particular tend to become very large as they devour existing physical-world-social networks (i.e., everyone can find their friends there), or they tend to fail quickly. A social platform can also collapse midstream via cascade processes as users depart – this is what happened to the early social network Friendster (Garcia, Mavrodiev, and Schweitzer 2013). Of course, every company wants to grow, but the incentives are much stronger for growth in platform companies than for other types of business, which have other ways to be profitable and aren't in danger of such failure cascades.

[9] Indeed, to some extent, code-sharing platforms have already evidenced some of the characteristic problems of social media, such as quasi-viral content that turns out to be malicious, in the form of security holes introduced by bad actors into popular open source libraries (e.g., Goodin 2018).

[10] Meta is perhaps the most famous example of this imperative for growth, which has been the subject of frequent discussion in the press (e.g., Roose, Isaac, and Frenkel 2020; Kantrowitz 2019; Hao 2021). This is doubtless because of Meta's particularly prominent acquisitions of Instagram and WhatsApp, which – especially the latter – are most understandable as investments in new groups of active users.

In practice this may also imply concentration because using platforms is costly (if only in time): If I have to search Amazon and Walmart and eBay and Mercari every time I shop for a product, I'm likely to want to abandon the smaller of these platforms as soon as practicable in favor of a platform that is more likely to have everything I need in one place.[11] This may imply a winner-take-all platform economy even aside from any anticompetitive behavior by companies, although companies may also promote growth in part to fend off competitors by ensuring that users who are economizing on on-platform time to meet some need which scales with addressable network size can meet it on that company's platform rather than a competitor's.

### 1.1.4 *Recommender Algorithms*

Positive network externalities have a downside for users: search costs. There is a point of diminishing returns to network expansion in which those costs overwhelm the benefits of additional other users for a given user. In order to prevent this potentially growth-limiting eventuality, platforms have strong incentives to invest in tools that facilitate their users discovering whatever they happen to be looking for. This is the economic logic behind the recommender algorithms that drive phenomena like Facebook's news feed. Recommender algorithms promote both scale (by reducing the costs to users of that scale) and engagement (by positively reinforcing platform use).

Here's another way of thinking about the same problem: because users tend to be both producers and consumers of content in platforms, and successful matches are profitable (because of the skim revenue model), companies seek to provide free marketing for users in their role as producers of high-quality content, where quality is defined (problematically) in terms of the predicted desirability of that content to other users in their role as content consumers. Of course, these same algorithms are also used in paid marketing: The same identification of user interests that facilitates the placing of cat pictures on my social feeds also facilitates advertisers identifying that I am likely to be interested in buying cat products.

Recommender algorithms systems have played a key role in numerous platform governance challenges, in a variety of ways, including:

- The capacity of malicious external actors to attack those algorithms (Carley 2020, 372; Beningi, Joseph, and Carley 2019).
- The capacity of companies themselves to manipulate those algorithms for inappropriate purposes, such as associated with infamous (and highly controversial) allegations that Facebook employees de-prioritized conservative news sources in 2015 (Nunez 2016).

[11] I'm using a bunch of "may" weasel words here because I'm neither an economist nor an antitrust expert, and the nature of platform markets tends to generate debate among those who know much more economics than I do. This paragraph probably deserves an unusually large grain of salt. For a discussion of other sources of concentration from network externalities, see Martin-Laborda (2017).

- The data collection and analysis that companies must engage in to make such algorithms effective (which largely work by attempting to use machine learning to predict what users will be interested in, for various observable proxies for interest such as measures of engagement or sales), data itself which may be abused, as when Amazon has made use of third-party sales data to inform its own (competitive) product offerings (Mattioli 2020) or the infamous Cambridge Analytica scandal.
- The possibility of those algorithms to amplify pathological behavior purely as a matter of accident, as when people tend to engage with outrage-inducing content, thus fooling recommender systems into thinking they want to see more of it. For example, one leaked 2016 presentation from Facebook suggested that fully "64% of all extremist group joins are due to [Facebook's] recommendation tools" like "Groups You Should Join" and "Discover" (Horwitz and Seetharaman 2020).
- The related capacity of such algorithms to promote misinformation and polarization by meeting user demands for false or biased content. For example, one 2022 study found that YouTube's recommendation algorithm recommended more "big lie" videos about the US 2020 election to those who had already bought into Trump's election fraud narrative (Bisbee et al. 2022). As the authors suggest, this might actually be a kind of effective (albeit extremely socially harmful) design from the user and (short-term) company standpoint, since the algorithm was recommending content that those users were likely to want (Bisbee et al. 2022, 16–17).
- The danger that such algorithms may also exacerbate underlying social inequalities; for example, employment and housing recommender systems may unintentionally engage in illegal race and gender discrimination by making predictions about to whom an advertisement may be "relevant" (Ali et al. 2019).

At the same time, we can conceptualize the content moderation side of platform governance as an outgrowth of the idea of a recommender algorithm. That is, if I'm on (say) eBay or Facebook, not only do I as a user want to be able to easily find the products or posts that I'm interested in buying/seeing without having to sort through too many unwanted products or baby pictures, I also want to make sure that the products and posts that I consume are, for example, genuine rather than counterfeit, honest rather than Russian-created misinformation, and the like. In that sense, we can understand content moderation aimed at decreasing the likelihood of unwanted discoveries as no different from recommenders to increase the likelihood of wanted discoveries, in the context of limited user attention. At a sufficient level of abstraction, both recommendation algorithms and content moderation reduce to the notion of a feed ranking (as in the ordering of items on Facebook's news feed or in Google or Amazon search results) in which each item–user pair has a score predicting that user's interest and how much the platform will benefit from displaying

that item to the user, and in which users are first shown highly scored items. To moderate is to reduce that score, to recommend is to increase it.

Recommender algorithms also tend to promote a kind of winner-take-all character among users producing content or selling products: Recommender algorithms tend to reward successful attention-getting with more attention-getting.

### 1.1.5 *Virality (with a Caveat)*

One consequence of recommender algorithms that is distinctive to social media platforms is the phenomenon of virality, in which certain kinds of content – particularly emotionally arousing content (Berger and Milkman 2012; Brady, Gantman, and Van Bavel 2019; Guadagno et al. 2013) – tend toward extremely rapid dissemination, and hence extremely rapid and large impact. Viral content entails a number of problems. First, as an operational matter, it can outpace the efforts of content moderators to control it. Second, the phenomenon also serves as an attack surface for hostile actors, who may seek to manipulate broader publics by producing such content and seeding it widely (Bowers and Zittrain 2020, 4). Third, because such content is often of low quality (e.g., fraudulently anger or outrage-inducing), it represents a kind of overall pollution of the discursive environment of such platforms.

Although virality is characteristic of social media platforms in which content can be reshared between users – particularly at low-cost and unreflectively, doubtless drawing on what Kahneman (2013) calls "system 1" thinking – it seems to me that some variation of it is, in fact, characteristic of all recommender system-driven arrangements that rely on pre-existing user data, and thus can be characteristic of transactional platforms, hybrid platforms such as Google search, and social media without a strong resharing affordance like YouTube. Where consumption of some content or product (including, e.g., some cheap drop-shipped garbage on Amazon) is of low cost, where it is by contrast more costly to acquire information about the quality of the content or product (i.e., to exercise the psychological restraint to check on the truth of the absurd claim one is retweeting or watching on YouTube, or investigate the origins of the junk one is buying), and where a platform uses consumption as a basis for recommendation, it is possible for low-cost consumption choices to become magnified. Those low-cost consumption choices are likely to be characteristically low quality insofar as low quality is correlated with low cost, as when it is both cognitively easier and worse to consume and share emotionalizing content, or when worse products are cheaper.

I mention this because it's easy to get caught up in the micro-affordances of platforms. For example, one idea with an immediate appeal for tamping down virality is to tamp down resharing – either by removing that functionality from platforms that have it, or by impeding its use (as with Twitter's occasional interstitial that pops up on shares of news articles asking one if one wants to read the article before retweeting it). Yet in a social media world no longer dominated by chronological feeds,

intervening in user behavior is only an indirect remedy for virality, for what matters is not what users choose to reshare but what the platforms choose to display. Twitter could, for example, simply reduce the probability of unread retweets appearing on one's news feed rather than scolding its users. The discourse around virality, like that on Twitter's attempts to reduce it, seems to implicitly assume a direct causal relationship between user actions and platform outcomes, one that conceals the role of product design in that relationship.

This is an important cautionary point with respect to the notion of platform governance as a whole: The insight, associated with scholars such as Klonick and Gillespie, that governance is part of the product platforms offer needs to be matched with the equally true point that product design is part of governance. The consequences of user behavior on platforms can be altered either by interventions on that behavior or by interventions on the impact of that behavior on other users.[12]

We might go further and suggest that the notion that user behavior has a direct effect on platform outcomes is a kind of ideology, in something like a (loosely speaking) Marxist sense, that is, a false belief about the social world and its fixed and necessary characteristics which supports the interests of those who hold power and derives from those relations of power.[13] Like an ideology, it seems to pose difficulties even for otherwise sound accounts of the existing system. Just to pick a single example, almost at random, in an otherwise excellent and insightful 2018 article defending a multistakeholder conception of platform governance, Natali Helberger, Jo Pierson, and Thomas Poell (2018) state that social media platforms "only partly control content distribution" (p. 2) and "content on social platforms only spreads if many users share it" (p. 7). Both claims are false: Social media platforms completely control content distribution and could easily bring it about that content spreads without user sharing.[14]

But after abandoning the ideology that *anything* a user does on a platform has an unmediated effect on any other user, the reader may wonder why we ought to have a category of "platform governance" focused on user behavior at all. To my mind, the fundamental constraints leading to a focus on user behavior are economic: A key element of the economic value of platforms to their users is that user behavior is conferred some causal role on other users' experience – PageRank was a huge part of the value of Google's early search; the fact that one's news feed contains the cat photos one's friends have chosen to share is a huge part of the value of Facebook – and

[12] Cf. Srivastava (2021, 4) who aptly categorizes functionality like the arrangement of Facebook's news feed as a form of governance – "herding" – alongside content moderation.

[13] This is a loose version of a combination of the first and second versions of Marx's account of ideology given by György Márkus (1983). We might also think of this as a kind of reification, but Marxist social epistemology is far too deep a sea for this book to explore.

[14] For example, in 2016, the current wave of debate about alleged social media political censorship was sparked by allegations that Facebook's "trending news" feature was actively edited to suppress conservative stories (Herrman and Isaac 2016).

these economic reasons (as well as, of course, the costs of more manual curation as opposed to making use of the economic value of the "labor" of users) impose a soft constraint on the extent to which what some users are presented with on platforms and their associated external effects can simply be handled by direct intervention on platform outputs as opposed to user inputs.

### 1.1.6 *Diversity of People and of Behavior*

The combination of the foregoing platform features leads to a vast diversity of people and thereby of behavior. Large-scale platforms involve people with lots of different interests and resources, and hence who are predisposed to behave differently from one another. And because of the network affordances such platforms offer, many of those people find themselves interacting with others whose interests and resources are dissimilar from their own, who may have different assumed cultural starting points for interaction, and who are otherwise importantly diverse.[15] This, in turn, leads both to behavioral novelty when those diverse interactors discover shared behaviors compatible with their interests, resources, cultural backgrounds, and so forth, and to conflict, when they do not.

This is a kind of converse of the niche-finding property, and one of the most interesting features of platforms is that they can go together. However, this is not wholly novel – it's also true of, for example, the city – in which the sheer scale and density of the urban environment mean that one can both find one's niche interests (because there are enough people to economically support serving them in a big city) *and* that when one is out and about on the street one is liable to be exposed to people who are different from oneself, and perhaps learn from one another and invent new ways of being in the world together.[16]

We should distinguish behavioral diversity and user diversity. User diversity is a consequence of scale. Behavioral diversity, however, is both a consequence of user diversity in the context of recommender algorithms promoting niche-finding and the simultaneous fortuitous interactions between different users who would never interact in other contexts. But they're also a feature of certain kinds of platform designs, themselves a consequence of the incentive to promote engagement.

Thus, the interactive affordances of platforms tend to be open-ended within the boundaries of the technical forms available to a company: YouTube doesn't exist to distribute specific kinds of videos; leaving aside concrete law or policy-based restrictions, it exists to distribute whatever videos its users see fit; the same can be said about eBay and goods.[17] The economic incentives for scale and engagement suggest

[15] Moreover, users seeking a market for their unique offerings (whether content or products) are, assuming an effective recommender algorithm, much more likely to find that in a larger network – hence there is probably an incentive for users with niche offerings to flock to larger networks.

[16] On the relationship between density and diversity, see, for example, Moroni (2016).

[17] Sometimes this is called "generativity" (Gorwa 2019b, 3).

https://doi.org/10.1017/9781108975438.002 Published online by Cambridge University Press

that a high degree of adaptivity to diverse user needs will be rewarded – and this further suggests that one key economic incentive of platforms will be to diversify the technical forms available in order to make additional types of interactions available to their users.[18]

The foregoing implications taken together lead to a high likelihood of interactions on multilateral platforms that tend to be unanticipated and unanticipable in advance. Content goes unpredictably viral. Emergent properties of densely networked interaction at scale crop up, seemingly out of nowhere. The broader ecosystem of content on such networks can be subject to high-speed evolutionary landscapes where particular kinds of activity propagate and disappear. The phenomenon of "context collapse" upsets people's ordinary intuitions about the audiences for their activity and hence leads to surprise interpersonal responses even on an individual basis (Marwick and boyd 2011) and then that scales too, and you end up with, well, Twitter.

## 1.2 PLATFORM PROBLEMS ARE LIKE STATE PROBLEMS: THERE ARE LESSONS TO BE BORROWED

In consequence of the above, platforms tend to pose severe challenges for the control of malicious or unintentionally harmful behavior. Novel kinds of behavior emerging at high speed among a vast and diverse populace in the context of positive feedback loops on attention make it difficult to identify the harmful behavior likely to appear, and difficult to enforce rules against identified types of harmful behavior.[19]

Those properties, in other words, generate distinctive governance difficulties that have been observed worldwide in recent years on social media. For example, the phenomenon of viral hoaxes has been particularly hard for such platforms to control not merely because there are internal and external conflicts about the role of such platforms in policing truth (although this is certainly the case), but also because such hoaxes can exist in a huge number of languages, in swiftly evolving idiomatic form, distributed by a huge number of users in a huge number of ways – and thus

[18] In particular, this incentive seems to drive at least some part of both Google's and Facebook's product strategies, such as with Google's persistent experimental creation (and routine destruction) of new ancillary services, like Google Reader and Google Flights, as well as acquisition of companies occupying different roles in interactive contexts, like Blogger, as well as Meta's acquisitions of companies specializing in different technical forms, such as virtual reality, and the expansion into new forms of interaction such as trade ("Facebook Marketplace") and dating. Schwarz (2017, 384) plausibly suggests that this tendency also arises from the monetization of user data, as it creates new profit opportunities derived from merging data across novel domains of user activity.

[19] Platforms attempt to leverage the same artificial intelligence techniques driving their recommender systems to also identify problematic behavior at scale and speed. See, for example, Mark Zuckerberg testimony before Senate Commerce and Judiciary Committees, April 10, 2018, available at www.washingtonpost.com/news/the-switch/wp/2018/04/10/transcript-of-mark-zuckerbergs-senate-hearing/ ("Now, increasingly, we're developing A.I. tools that can identify certain classes of bad activity proactively and flag it for our team at Facebook").

https://doi.org/10.1017/9781108975438.002 Published online by Cambridge University Press

because any platform wishing to control those hoaxes would require the capacity to identify their distribution in these ever-changing forms at speed and scale – a gigantic problem. Moreover, those platforms must be capable of developing policies that describe forbidden behavior at a sufficient level of generality to be administrable without making countless costly and inconsistent one-off discretionary decisions, but which are still capable of being sufficiently tailored to individual situations to achieve the underlying goals of those policies.

In these respects, platforms bear an important functional isomorphism to key forms of political organization such as cities and states. Polities are also, on some important ways of thinking about the role of political organization, fundamentally infrastructural – polities provide a broad array of affordances, such as physical security and enforceable contracts, in order to permit people to engage in diverse behavior at scale. This is a notion that's shot through legal and political theory, or at least the small-l liberal kind – if there is a foundational proposition of liberalism, it is that the function and justification of the state is to help people to engage in private projects. For example, some scholars think that a key function of law is to solve coordination problems getting in the way of people achieving their collective well-being – to do stuff like enable people to decide which side of the road we all drive on (e.g., Fuller 1969; Postema 1982, 182–86). A related way of thinking about the function of law and of the state is "determination," identifying principles of (property, etc.) law to be enforced in order that people may have secure access to means by which to pursue their private ends (e.g., Ripstein 2004, 11–13, 26–29; Beitz 2018, 425–27). Contract law is often said to be explained and justified in view of the fact that it permits people to make complex, binding, plans relying on one another, and thus solve all kinds of problems of trust and credible commitment that would otherwise hinder socially and individually useful joint projects among individuals (e.g., Dagan and Heller 2020). Political philosophers can recognize this idea in Rawls's supposition that the original position models a decision-making process motivated by the goal of framing a society in which its members can develop and pursue their conceptions of the good – and that the products of that society are divided as the "benefits … of social cooperation" (Rawls 1999, 4).

In that sense, one could argue that *the state itself is a platform*.[20] Moreover, the interests of states align with the interests of their citizens in the same broad way that the interests of platforms align with the interests of their users. States typically derive

[20] Some in the technology industry have offered a version of this suggestion. For example, software industry publisher Tim O'Reilly (2011) analogized government to hardware platforms insofar as they offer infrastructure for collective action among their people. Arguably, instead of the state, we might conceptualize the law itself, in contexts including, but not limited to, states (e.g., Ellickson 1991) as a platform, though I cannot explore this idea here. It is interesting to note that law has positive network externalities to the extent that the transaction costs of far-flung and diverse economic and social actors are reduced by having a common law to which they may appeal – a dynamic that contributes to some economic historians' accounts of the *lex mercatoria* in medieval Europe (Milgrom, North, and Weingast 1990; but see Kadens (2012, 2015) for historical dispute as to its existence).

revenue from taxing productive activity; one influential theory of how states learn to refrain from recklessly expropriating and abusing their citizens observes that to the extent rulers can extract rents from that activity, they do better to provide their people a minimum of security in order to facilitate the activity's occurrence (Olson 1993).[21] At the highest level of abstraction, advertising revenues on Meta or commissions on eBay are a kind of a tax on user activity.

We might further say that just as moderation is part of the product of companies, the rules of property and contract and the like are part of (or the fundamental) product of governments. In fact, there have been cases where governments or rulers have directly sold access to the legal form of resolving disputes. For example, medieval English kings received meaningful revenue from the sale of adjudication, either in the form of court fees or in the form of kickbacks from adjudicating cases promptly (Ramsay 1925, 1:1–2, 58; MacPhail 1926, 251). In the contemporary world, we can think of paid private arbitrators as the market supply of adjudication undersupplied by the state. Contemporary court fees also represent, in some sense, the idea of legal resolution having a monetary value that the state can partially exploit.[22] Moreover, contemporary choice of law and choice of forum rules and their use in the commercial world might represent the idea that states compete in their supply of legal rules and adjudication.[23]

Similar points are true of sub-state forms of political organization. For example, a city, understood as a collection of affordances for social interaction analyzed by influential urban theorists spanning a range from Jane Jacobs to Richard Florida, resembles one of the multilateral platforms described above. The success of a city is a consequence of its ability to promote beneficial social and economic activity by bringing together, at scale, individuals seeking to participate in that activity with one another, and cities provide affordances in the form of physical spaces such as roads and parks, as well as incentives, such as economic development grants, in order to make these interactions possible; cities also fund their activities by taxes qua skim off the top of such activity. The density (which can be represented as the potential for each resident to have more network connections with others due to sheer physical proximity) of a city itself provides some of its key advantages, both in terms of practical economies of scale and in terms of positive network externalities from, for example, locating sellers and buyers and workers and jobs in the same place. Cities also tend to promote diversity and novel forms of behavior and, with that, novel

[21] Contemporary officials in modern democratic states, who do not personally receive a share of taxes on productive activity, still have an incentive to facilitate that activity (a.k.a. "economic growth") to the extent that electoral advantages to prosperity substitute for the direct receipt of tax revenues by individual leaders.

[22] Some American municipalities use court fees charged to those who encounter the criminal justice system as a source of revenue (Sances and You 2017; Mughan 2021). The US federal courts also have sold access to court records at a price which critics have alleged exceeds the cost of operating the system (DeWitt 2010).

[23] Though this competition is sometimes, perhaps even usually, pernicious (e.g., Slemrod and Wilson 2009; Anderson 2015).

enforcement and safety challenges – plus the same density that helps sellers find buyers also helps criminals find victims.[24]

Unsurprisingly, given that they have (at a certain level of abstraction) similar incentives and serve similar activity-enabling functions in the lives of their people, polities have experienced many of the same governance challenges with which platforms now struggle. For example, perhaps the most fundamental challenge of using the form of law to govern human behavior is to strike the balance between general rules to be consistently enforced and the capacity to adapt to specific situations in the universe of potentially infinite human behavioral novelty (e.g., Schauer 1991). This is a problem for polities and platforms alike. In the abstract, the balance between general rules and adaptability to specific situations in application is a problem associated with *all kinds of behavioral regulation* – in principle, for example, it applies to situations as prosaic as regulation by parents or employers (how to set attendance rules at work that are both fair to all, yet accommodate individuals' special circumstances?).

Theorists in the domain of polities have wrestled with these problems for a long time. The management of open-ended interaction affordances in the context of scale and diversity has been a central concern for political theory as long as that discipline has existed. In Ancient Greece, for example, Aristotle famously wrestled with the problem of scale, suggesting that a polity ought to be small enough that its members know one another's characters, in order that they might effectively select officials and conduct litigation.[25] Contemporary scholars have identified that scale and diversity entail distinctive problems of central state planning in that central policymakers will often lack knowledge to be found only on the edges of the community; important academic work focusing on this problem has ranged from identifying it as the source of disaster in modern governance (J. C. Scott 2008) to favoring markets as a solution to information problems (Hayek 1945), to explaining the success of democratic institutions for aggregating dispersed information (Ober 2008) (much more on all of these in Chapter 3).

[24] Physical economic markets also have something of this platform character. Economic historians have written about institutions such as the Champagne Trade Fair, a physical site (or set of them) – supported by law but having a physical importance in its capacity to bring trading partners together. Consider also the New York Stock exchange and more broadly the contemporary electronic stock trading system, qua infrastructure for making trades, even novel trades that can be adapted to functional forms afforded by the system. Such markets again seem to partake of positive returns to scale, with returns that are internalized by market operators and those that are bestowed on participants. Contemporary financial markets also seem to partake of phenomena that bear at least a superficial resemblance to virality. For example, it may be sensible (I am not sure) to represent feedback effects associated with "flash crashes" as a kind of viral propagation of toxic false information about some market participants, represented in the prices associated with such participants (on flash crashes, see, generally, Lange, Lenglet, and Seyfert 2016). Of course, this recently merged with the problem of social media as Reddit created a kind of "flash bubble" in stocks like GameStop. Such markets also produce beneficial as well as pathological novelty, including, for example, the exotic financial instruments that were used to conceal the riskiness of subprime mortgage loans in the mid-aughts. Maybe some of the policy proposals described in this book could also apply to stock exchanges, though I don't dare try to defend such a notion here.

[25] Aristotle, *Politics*, book VII, 1325b–1326b.

https://doi.org/10.1017/9781108975438.002 Published online by Cambridge University Press

The benefits and challenges of diversity at scale are an axis around which debate has revolved in areas as widespread as the provision of social welfare benefits (Kymlicka 2015), the creation of patriotic identities (Müller 2007b), and novelty-driven social progress and economic growth (Thomas and Darnton 2006). One fair description of the first premise of contemporary liberal political philosophy is that its goal is to describe political institutions that permit a large population that is diverse in both interests and values to live together while maximizing the freedom of each individual to pursue their individual interests and values. This is, for example, John Rawls's (1985, 233) conception of the "two moral powers" of each citizen. Almost every public good provided by a political institution can be described as a set of affordances to pursue the development and exercise of those moral powers. On the opposite side of modern liberal theory, Robert Nozick (1974) implicitly applied the idea of positive network externalities and returns to scale in his story of how a libertarian state might arise from a "dominant protective association."[26]

Given the broad array of functional similarities between internet platforms and polities, it should be unsurprising that ideas from liberal political philosophy have leaked into the self-understandings of platform users as well as the companies themselves. The most salient example is that of the idea of free speech – there is persistent controversy in the United States about the extent to which First Amendment rights ought to apply against decisions of those platforms.[27] At various points in their corporate histories, social media companies have proudly trumpeted their free speech credentials (Gillespie 2010, 356), with Twitter famously describing itself as "the free speech wing of the free speech party" (Halliday 2012); and positive accounts about the impact of social media have focused (at least prior to 2016) on its capacity to facilitate core First Amendment-type values such as politically dissident speech in autocracies (e.g., Passini 2012).

Nor are those normative claims necessarily as misguided as they might seem from the perspective of a frustrated 2023 civil rights lawyer sick of explaining to people why there's no First Amendment right to spout racial slurs on Facebook.

[26] Early in the internet age, David Post (1995) borrowed from Nozick's idea of competition between providers of rules to sketch out a vision for the development of governance of online activity.

[27] The simplest answer in terms of existing legal doctrine (which is more of an "is" than an "ought" issue) is "not at all," although there is an important, although largely disused, corner of First Amendment doctrine according to which the operators of the "functional equivalent of the traditional public square," such as owners of company towns, may be treated as quasi-governmental actors subject to the First Amendment. See, for example, *Marsh v. Alabama*, 326 U.S. 501 (1946); *Robins v. Pruneyard Shopping Center*, 23 Cal.3d 899 (1979). Scholars have analogized social media both to cities in the form of the town square and to governments in the form of speech regulators in order to consider the possibility that such doctrines may be applied to Facebook and Twitter (e.g., Tussey 2014). There has also been legislation reflecting this position passed in both Texas and Florida, and as of this writing there are ongoing controversies about the constitutionality of those statutes (which controversies are, blessedly, beyond the scope of this book).

Scholars have recognized that "the state" is not an organic entity that can unproblematically be defined from outside – as Bevir and Rhodes (2015) explain, the state is a kind of sociological and narrative phenomenon (in their words, "cultural practice") derived from the self-understandings of participants in activities of governing. Similarly, international relations scholar Swati Srivastava (2022) has defended a theory of "hybrid sovereignty" according to which private organizations ranging from the English East India Company and Blackwater's mercenaries to nonprofits like Amnesty International have become entangled with allegedly core aspects of state sovereignty or even have such aspects delegated to them.[28] Blackwater, for example, carried out combat in Iraq on behalf of the US government, defying the notion that the use of military force is the absolute heart of modern indivisible state sovereignty. Platforms do something similar. The intermediary liability framework of statutes like the DMCA, for the most obvious example, seems to recruit platforms to carry out copyright adjudication and enforcement on behalf of the state in a manner not terribly dissimilar from how Blackwater was recruited to carry out war, except that the tool used to do so was the risk of liability rather than contract (i.e., a stick rather than a carrot).[29] Human rights scholar Molly Land (2019, 404–8) has already argued that these intermediary liability frameworks are a form of delegated state authority for some international law purposes.

In my view, there's no a priori reason to rule out platform companies from being part of such a state practice or hybrid sovereignty, particularly when their personnel are using familiar forms of law-like reasoning to issue authoritative determinations about individual behavior in their limited contexts – and hence no a priori reason to deny that platforms are subject to the normative standards we apply to states, although our judgment about those normative standards should not be taken to determine the separate (though obviously entangled) question as to whether legal standards such as the US First Amendment should apply to them. This book ultimately takes no position on whether platforms ought to be treated as carrying out some form of state action or manifesting hybrid sovereignty, which is not a problem that needs to be resolved for the argument to go through. But I *do* contend that it should be clear, perhaps even uncontestable, that platforms do carry out some tasks that are meaningfully similar, in both their social role (e.g., managing political discourse) and in the constraints they face (i.e., managing scale and diversity), to tasks carried out by states.

[28] For a much more sinister example, I have elsewhere suggested that white supremacist paramilitary organizations had, in some legally and theoretically relevant senses, become part of or taken over part of the state in at least some time periods in the US South after the end of Reconstruction (Gowder 2021, 92–93; 2018a, 348–51).

[29] It may be that intermediary liability can serve as a gloss on the form of hybrid sovereignty which Srivastava calls "contractual hybridity." It should perhaps be unsurprising that Srivastava (2021) has also written one of the few articles from the institutional side of political science offering serious attention to the notion of platform governance as governance.

Those observations suggest that efforts, whether governmental or private, to build institutions capable of shaping aggregate behavioral outcomes on networked internet platforms can benefit from drawing on our existing knowledge (theoretical and practical) of political and legal institutions.[30] Indeed, such efforts are already ongoing – it is not a coincidence that organizations such as Facebook and Google, for example, recruit lawyers – members of the profession built to handle precisely such governance problems – to head their content policy teams,[31] or develop organizational forms explicitly derived from governments.[32] Nor is it coincidental that scholars such as Kate Klonick (2018, 1641–48) have explicitly analogized these efforts to the acts of government policymakers and identified that the process of actually implementing platform rules looks remarkably like administering a legal system. It so happens that the functional forms of legal systems developed as responses to problems remarkably like those which platforms face.

For those reasons, it should now be clearer why I contend that theoretical approaches to political and legal institutional analysis may help us understand the problems of networked internet platforms. Moreover, I shall suggest that the institutional tools of polities, such as the rule of law and techniques associated with ecological governance and urban design, may help us alleviate them.

[30] Cf. Gorwa (2019a, 859–60), observing a shortage of input from political science in platform governance.

[31] For example, former federal prosecutor Monika Bickert leads a team making content policy at Facebook (Zuylen-Wood 2019) (for purposes of full disclosure: Bickert attended Harvard Law School at the same time as I did, though we haven't spoken in many years). Google's head of "Trust and Safety" is not a lawyer, but is a White House veteran (Canegallo 2019). Other high-ranking personnel in Google's/Alphabet's content policy operation have been lawyers; for example, a former ACLU lawyer occupied a number of leadership roles in YouTube content policy. (See the LinkedIn Profile of Juniper Downs, www.linkedin.com/in/juniper-downs-5245879/, visited April 20, 2020.) According to a profile by the India Times, Twitter's pre-Musk general counsel bore responsibility for its content policy (India Times 2020).

[32] For example, Meta's content moderation oversight board was conceived as a kind of "Facebook Supreme Court" expressed in exactly those terms (Ingram 2019).